Sentencing Council

Crown Court Sentencing Survey

Record Level Data

User Guide

Coverage:	England and Wales
Period:	1 January 2011 to 31 December 2011
	1 January 2012 to 31 March 2012
	1 April 2012 to 31 December 2012
	1 January 2013 to 31 December 2013
	1 January 2014 to 31 December 2014

Contact details

For further information or to inform us of your research, please contact the Office of the Sentencing Council at research@sentencingcouncil.gsi.gov.uk

More information on the Crown Court Sentencing Survey is available at https://www.sentencingcouncil.org.uk/analysis-and-research/crown-court-sentencing-survey/

Version number: 1.0

Issue Date: 25 June 2015

Contents

1 Introduction 3

2 Overview of the Crown Court Sentencing Survey 4

3 Time period of the databases 4

4 Structure of data files in the databases 5

5 Content of datasets 6

6 Matching files 7

7 Conventions for missing data 7

8 Data quality and interpretation 8

Annex 9

1 Introduction

This document contains important information on the use and interpretation of the Crown Court Sentencing Survey (CCSS) record level data for 2011 to 2014. Following the publication of the latest CCSS results on 25 June 2015, anonymised record level data for 2014 has been released. Data for 2011, 2012 and 2013 was published previously in February 2015 but has also been included for completeness.

Before you make use of these datasets, it is recommended that you read the information provided in this document, familiarise yourself with the layout of the survey forms and refer to the accompanying metadata document.[1]

A survey form was expected to be completed for every new criminal case sentenced at the Crown Court in 2014. Where an offender was being sentenced for more than one offence on the same indictment, the sentencing judge was required to consider only the most severe or "principal" offence.

The datasets are provided in csv format and contain formatted values rather than coded information to make it easier for users to understand what the data contain. The datasets relate to adults only; records relating to offenders that were under the age of 18 at the time of sentencing have been removed due to the risk of these individuals being identified from the data.

[1] Copies of the survey forms as well as the metadata document can be accessed on the Council's website at https://www.sentencingcouncil.org.uk/analysis-and-research/crown-court-sentencing-survey/

2 Overview of the Crown Court Sentencing Survey

The Crown Court Sentencing Survey (CCSS) began on 1 October 2010. Following a review last year, the Sentencing Council decided to stop gathering data using the CCSS, and data collection ceased on 31 March 2015. This is therefore the final publication of record level data from the survey.

In 2014, the survey data were collected using twelve different offence form types.[2] The data has been used to produce an annual publication which provides a high level summary of the data collected. The latest publication is available at http://www.sentencingcouncil.org.uk/analysis-and-research/ccss-annual-2014-results/.

The survey has helped the Sentencing Council to develop new guidelines, to make sure existing guidelines are working as intended and to inform the wider public about how sentencing decisions for offenders are made. In the future, the Council will conduct bespoke data collection in both the Crown Court and magistrates' courts for specific guidelines, thereby extending analysis of sentencing practice into magistrates' courts.

The Council hopes that publishing this record level data will enable others to make use of the data to conduct further research on sentencing and sentencing practice.

We are happy to provide advice on the use and interpretation of this database. If you are making use of these datasets and have any queries, you are welcome to get in touch using the contact details provided at the front of this guide.

3 Time period of the databases

The record level data consist of five separate databases covering 2011 to 2014. They have been split so that some structural discontinuities in 2012 are highlighted. It is important to understand the reasoning behind this split so that if the individual datasets are amalgamated, for example for time series analysis, they are used appropriately.

The main structural discontinuity occurs from April 2012, when a new organisation took over the responsibility of processing the survey forms and transferring the information into a database. The physical layout of the forms was completely redesigned, resulting in an improvement in the quality of the data collected. However, the structure of the dataset delivered was very different and incompatible with that produced previously. As a result the data delivered prior to April 2012 was converted to a format as consistent as possible with that delivered by the new processors. However, the conversion process could not be carried out in a way that preserved all the information. As a result there is a significant structural break in the data at the beginning of the second quarter of 2012 (1 April 2012) and users should be aware that there are some inconsistencies with the data delivered before then. These inconsistencies are discussed in the annex.

The other structural discontinuity to note relates to the instances where a survey form was not completed when it should have been, henceforth known as "unit" non-response. Data for 2012 onwards[3] has been

[2] In 2014, following the introduction of the new definitive guideline for fraud, bribery and money laundering offences and the definitive guideline for sexual offences, fraud offences were removed from the theft, dishonesty and fraud form and completed on the new fraud, bribery and money laundering form, and the sexual offences form was replaced by two new forms; sexual offences and indecent photographs of children.

[3] The information is not available for 2011 data.

enhanced with information from the Ministry of Justice Court Proceedings Database, CPD.[4] This enhanced data should assist users with obtaining national estimates better aligned to those in the CCSS published outputs. These records can be identified using the 'CCSS_FORM_COMPLETED' variable. They can also be identified by the fact that they only contain information for the key variables: sentence month, age, gender, form type, type of offence, sentence outcome and custodial sentence length.

These structural differences in the databases are summarised in table 1.

Table 1: Structural differences in the CCSS record level databases

Database	Period	Data properties
2011	January to December 2011	Original data converted to new structure consistent with the data delivered since April 2012. Information on unit non-response not included.
2012, quarter 1	January to March 2012	Original data converted to new structure consistent with the data delivered since April 2012. Information on unit non-response included.
2012, quarters 2 to 4 2013 2014	April to December 2012 January to December 2013 January to December 2014	New data structure. Information on unit non-response included.

4 Structure of data files in the databases

Each database consists of several CSV files of related content and similar structure; one overall or **main** file and several **additional** form-specific files.

Each **main file** contains all records during the relevant period but only includes information that is common across all offence form types such as offender details, sentence outcome and guilty plea information. This file should be used when analysis across all offences is required.

The **additional files** comprise a separate file for each form type covering the specified period. The files contain information already included in the main file but in addition, they also contain information that is specific to that form, for example the particular aggravating and mitigating factors applicable to that form. Note that some of the additional files only contain information for records completed on the **correct** survey form (as assessed by the type of offence).[5] This is to allow robust analysis to be obtained. Information on the number of records in each file is available in the annex.

[4] A database maintained by the Ministry of Justice, of all principal offences sentenced at the Crown Court and used to produce the MoJ quarterly criminal justice statistics publication.

[5] This is only relevant for the databases relating to 2011 and 2012 as improved data processing methods have been employed since 2013. The selection of aggravating and mitigating factors vary across all different form types, reflecting the differing nature of

This structure of the data files has been adopted so as to streamline the amount of variables present in each file, thereby reducing their size and improving processing speeds. A large amount of information is collected on the survey forms but most is only applicable to specific offence form types. If a single dataset were provided, it would consist of a large number of variables and a large amount of redundant data where information was not relevant.

Note that information on the level of seriousness has not been included in the main file but has instead been made available in the additional files only.[6] This is because it is not appropriate to include this information in a single file containing the full range of offences, because you cannot compare the seriousness of a shop-lifting offence, for example, to that of a grievous bodily harm offence. It is also not possible to compare the level of seriousness between offences within the same form, for example you cannot compare the seriousness of a domestic burglary offence to that of an aggravated burglary offence.

5 Content of datasets

The accompanying metadata Excel spreadsheet[7] contains details of the specific variables included in each file. You should refer to this metadata document before and during analysis of the database.

The variables present in the main file and the additional files are shown in table 2.

Table 2: List of variables available in files

File	Variables
Main	Unique identifier Month of sentence Age of offender Offender gender Form type Flag for whether a CCSS form was completed Flag for whether the correct form was used[8] Type of offence Sentence outcome Length (immediate custodial sentences only) Number of previous convictions taken into account Whether guilt was indicated at a police station Stage guilty plea was entered Guilty plea reduction that was given Flag for whether the guilty plea was entered at the first reasonable opportunity
Additional files	As above but in addition: Level or category of offence (for offences that have a sentencing guideline) Aggravating factors Mitigating factors

the offence types covered. If details of an offence are completed on the wrong form, then a consistent set of factors would not have been available for the judge to tick. This resulted in a small loss of records in each of the additional files in 2011 and 2012.

[6] This information has not been included for arson and criminal damage offences and driving offences because they do not currently have a sentencing guideline. The information is not recorded for drug offences sentenced under the definitive guideline as they follow a different model for assessing seriousness.

[7] Available at https://www.sentencingcouncil.org.uk/analysis-and-research/crown-court-sentencing-survey/record-level-data/

[8] Applicable to forms where a definitive guideline was introduced during 2014, i.e. fraud offences and sexual offences only.

6 Matching files

Depending on the type of analysis required, you may need to aggregate data spanning several files. A unique identifier is available in all files to enable you to do this more efficiently. This is the key variable for matching the main file and the additional files within a particular database and should be specified as the unique key when matching. All files on the database have been sorted by this variable so datasets should match easily using standard statistical software.

The unique number comprises of four parts which are separated with a hyphen (-):

1. An abbreviation that identifies whether the information was recorded on a form linked to a Sentencing Council ('sc') or Sentencing Guidelines Council ('sgc') guideline;

2. A three letter abbreviation for the form type that was completed (or should have been completed in the case of unit non-response for records since 2012). For example, 'apo' for assault and public order and 'bur' for burglary;

3. A 6 digit number used to identify the time period that the record relates to. The first 2 digits of the number represent the year (that is, 11 for 2011, 12 for 2012, 13 for 2013 and 14 for 2014), and the remaining 4 digits indicate the quarter included.

   1234 means all quarters 1 to 4
   0001 means only quarter 1
   0234 means quarters 2 to 4; and finally

4. A 6 digit number for the serial number.

An example of a unique identifier is 'sc-apo-120001-002287', which refers to an assault and public order sentence under new Sentencing Council guidelines during quarter 1 of 2012. The structure of the identifier should make it easier to identify specific records when you have amalgamated data across years and across offence form types.

7 Conventions for missing data

There are two types of missing data in the record level databases: item non response, where some information from a form has not been completed, and unit non-response when the form is not completed at all. The conventions for handling these are outlined below.

Item non-response

Most of the data provided are as completed by the judge at sentencing and corrections have not been made. In these cases, missing survey responses have been coded as "Not asked" where a question was not applicable and "Not answered" where the question should have been answered.

Corrections have been made to a few variables where missing information can be taken from the Ministry of Justice Court Proceedings Database (CPD). The variables that have been corrected in this way are offender age, offender gender, month of sentence, sentence outcome and custodial sentence length. Even after correction, there are still some custodial sentences with missing sentence length. These have been recorded with a dot (.) along with non-custodial sentences. Age at date of sentence is provided for all individual offenders but is recorded with a dot for companies.

Unit non-response

The 'CCSS_FORM_COMPLETED' variable indicates if a CCSS form had been submitted for the sentence passed; recorded as 'Yes' if recorded on the correct CCSS form, or 'Yes – incorrect form' if completed on the incorrect CCSS form. For 2012 data onwards, information for unit non-response (recorded as 'No') is also included within this variable.[9] This information is obtained from the CPD and for these unit non-response records, information is included for offender age, offender gender, month of sentence, sentence outcome and custodial sentence length. This additional information should assist users in obtaining national estimates.

Where there is unit non-response, limited data has been picked up from the CPD (see paragraph above). All other variables on these records are recorded with a dot.

8 Data quality and interpretation

The survey forms have been kept to a single page to minimise the burden placed on judges in responding to it. Therefore, although the forms capture a large part of the information taken into account when sentencing, they will not capture the full range of considerations made by a judge. Furthermore, the forms only indicate the presence of particular factors in a case - they do not record the relative "weight" of these factors on the final decision. These points are important to bear in mind when interpreting any findings from the record level data. Although trends may be identified, there may be reasons not captured by the survey that are the true cause.

Any data collection of this kind will be subject to some recording and processing error. Although a number of measures have been taken to clean and validate the survey data, there may still be some error in individual records. Revisions are made to the datasets when a significant error has been found which affects the publication. For further information refer to Annex B, published alongside the CCSS publication at http://www.sentencingcouncil.org.uk/analysis-and-research/ccss-annual-2014-results/.

Results produced from the record level datasets may vary slightly from the results presented in the respective CCSS annual publication. The reasons for this are outlined below:

- The results presented in the annual publication cover all sentences at the Crown Court, including those where the offender was under the age of 18 at the time of sentencing. The record level data only include information for adult offenders.
- Since completing the publications for 2011 and 2012, further validation routines have been developed and applied to the data.
- Item results presented in the 2012 annual publication have been "weighted" to correct for missing data.

Finally, over the period of data collection there have been some changes either to the forms or to the way the data collected has been processed. More information on this is available in the annex.

Feedback

We are keen to hear your views regarding the content, structure and ease of use of the record level databases and accompanying documentation.

Please contact us using the details at the beginning of this document.

[9] This information is not available for 2011 data.

Annex

Table A1: Number of records in each file

File	Database				
	2011	2012 Q1	2012 Q2 to Q4	2013	2014
Main	56,487	23,839	61,975	84,498	84,886
- of which on correct form	54,049	13,327	35,592	50,744	53,420
- of which on incorrect form	2,438	820	170	-	1,118
- of which no form was returned	-	9,692	26,213	33,754	30,348
Arson	896	191	649	814	827
Assault - new	5,420	3,094	8,553	11,510	12,786
Assault - old	9,081	181	-	-	-
Burglary - new	-	656	4,943	6,753	6,838
Burglary - old	6,679	967	-	-	-
Death	620	150	537	710	762
Driving	1,991	495	1,318	1,886	2,141
Drug - new	-	-	6,364	9,373	10,150
Drug - old	9,000	2,426	-	-	-
Other	4,504	990	3,017	4,545	5,238
Robbery	3,425	924	2,212	2,985	2,809
Sex (excluding indecent photographs of children) - new	-	-	-	-	2,059
Sex (indecent photographs of children) - new	-	-	-	-	768
Sex - old	3,553	957	1,906	3,875	1,905
Fraud	8,880	2,296	6,093	8,293	530
Theft	8,880	2,296	6,093	8,293	7,725

Changes to survey forms or data processing affecting the databases

There are five main changes that can impact on the data collected:

1. New guidelines
2. Changes to general form layout
3. Changes to existing questions
4. New questions
5. Change in processing data

Table A2 shows a log of the changes that you should be aware of if you carry out analysis over more than one database.

Table A2: Change log

Type of change	Database affected	Details and impact on database
New guideline	2011 2012 quarter 1	On 13 June 2011, the Sentencing Council definitive guideline for assault offences came into effect, replacing the previous Sentencing Guidelines Council guideline for assault offences. A revised version of the assault form was issued to replace the old form. This revised version differs significantly from the previous form to reflect the new structure of Sentencing Council definitive guidelines. There are therefore two files for assault available in the 2011 database, an "old" form file and a "new" form file. Also, due to delays in removing the old style forms from circulation, there are a few records in the 2012 quarter 1 database.
Change to general form layout	2012 quarters 2 to 4	In April 2012, the Sentencing Council redesigned all of the survey forms following feedback from judges on the ease of completion of the forms. However, no changes were made to the questions asked. The quality of data received from April 2012 is therefore higher than that collected previously.
New guideline	2012 quarter 1	When the definitive guideline for burglary offences was introduced on 16 January 2012, a new survey form was created. Prior to this, burglary offences were recorded on the theft, dishonesty, burglary and fraud form and have been recorded in the "old" burglary form file. There are therefore two files for burglary available in the quarter 1, 2012 database, an "old" form file and a "new" form file.
New guideline	2012 quarter 1 2012 quarters 2 to 4	Due to the introduction of a new definitive guideline for drug offences on 27 February 2012, the previous drugs form was replaced with a new form. However, as forms are circulated to court centres on a quarterly basis, new form types for drug offences were not introduced until April 2012. Therefore, unlike assault and burglary which have two file types in some databases during the transition to a new guideline, for drug offences there is only one file type present in each database. In quarter 1 2012, only the "old" forms are present while in the quarter 2 to 4, 2012 database onwards, only the "new" form files are present.
Change in data processing	Common assault records in Assault files for	Common assault is a summary only offence and is not usually tried at the Crown Court. However, if it is racially or religiously motivated (a section 29 offence) it is triable either way. It is also possible that the common assault cases may

Type of change	Database affected	Details and impact on database
	2011 2012 quarter 1 2012 quarters 2 to 4	well have been originally charged with a more serious offence, for instance ABH, but the change reduced before sentencing. The section 29 indicator has not been included in these datasets due to subsequent data processing issues. However, it has been included in the 2013 and 2014 databases.
Change to existing question	2012 quarters 2 to 4	The extra option 'Don't know' was included for the question on whether guilt was indicated at the police station. The options that can be ticked have increased from two to three with the introduction of this new option.
Change to existing question	2012 quarters 2 to 4 2013 2014	The response to the question on the guilty plea reduction given changed from a categorical to a continuous variable. However, court centres were asked to use up the old versions of the forms and therefore most of the responses to this question in the quarter 2 to 4, 2012 database were based on the categorical options present in the old forms. Therefore, any whole number responses were changed to the categorical options used in the previous databases. By 2013, most court centres had used up the old versions of the form and therefore, for the 2013 and 2014 databases, the discount has been recorded as a whole number where available.
Change in data processing	2012 quarters 2 to 4	For additional files that contain the level or category of offence variable, the additional option of "no guideline" has been included.
Change in data processing	2012 quarters 2 to 4	The proportion of records for which the question on number of previous convictions taken into account was not answered has reduced drastically due to improved data processing for this question.
Change to existing question	2012 quarters 2 to 4	The implementation of the Legal Aid Sentencing and Punishment of Offenders Act in 2012 changed the disposals available to judges when deciding on a sentence. This affected the options available on sentence outcome from December 2012.
New guideline	2014	When the definitive guideline for sexual offences was introduced on 1 April 2014, a revised version of the sexual offences form was issued to replace the old form. Because of the complexity of the sexual offences guideline, two new survey forms were introduced: a form for sexual offences excluding indecent photographs of children) and another for indecent photographs of children. There are therefore three files for sexual offences available in the 2014 database; the "old" sexual offences form file, the "new" sexual offences form file, and the "new" indecent photographs of children form file.

Type of change	Database affected	Details and impact on database
New guideline	2014	When the definitive guideline for fraud, bribery and money laundering offences was introduced on 1 October 2014, a new survey form was created. Prior to this, fraud offences were recorded on the theft, dishonesty and fraud form and have been recorded in the theft form file. There are therefore two files containing fraud offences available in the 2014 database; the theft form file and the "new" fraud, bribery and money laundering form file.